Available online at www.sciencedirect.com
SCIENCE DIRECT®
ACADEMIC PRESS
Cognitive Psychology 45 (2002) 375–412
Cognitive Psychology
www.academicpress.com

Is there a natural order for expressing semantic relations?

Lisa Gershkoff-Stowe[a,*] and Susan Goldin-Medow[b]

[a] Department Speech and Hearing Sciences, Indiana University, Bloomington, IN 47405, USA
[b] University of Chicago, Chicago, USA

Accepted 21 December 2001

Abstract

All languages rely to some extent on word order to signal relational information. Why? We address this question by exploring communicative and cognitive factors that could lead to a reliance on word order. In Study 1, adults were asked to describe scenes to another using their hands and not their mouths. The question was whether this home-made "language" would contain gesture sentences with consistent order. In addition, we asked whether reliance on order would be influenced by three communicative factors (whether the communication partner is permitted to give feedback; whether the information to be communicated is present in the context that recipient and gesturer share; whether the gesturer assumes the role of gesture receiver as well as gesture producer). We found that, not only was consistent ordering of semantic elements robust across the range of communication situations, but the same non-English order appeared in all contexts. Study 2 explored whether this non-English order is found only when a person attempts to share information with another. Adults were asked to reconstruct scenes in a non-communicative context using pictures drawn on transparencies. The adults picked up the pictures for their reconstructions in a consistent order, and that order was the same non-English order found in Study 1. Finding consistent ordering patterns in a non-communicative context suggests that word order is not driven solely by the demands of communicating information to another, but may reflect a more general property of human thought.

* Corresponding author. Fax: 1-412-268-2798.
E-mail addresses: gershkof@indiana.edu (L. Gershkoff-Stowe), sgm@uchicago.edu (S. Goldin-Medow).

0010-0285/02/$ - see front matter
PII: S0010-0285(02)00502-9

Keywords: Gesture; Communication; Language processing; Sentence structure

1. Introduction

> If I wish to communicate an intelligible idea about a farmer, a duckling, and the act of killing, it is not enough to state the linguistic symbols for these concrete ideas in an order, higgledy-piggledy, trusting that the hearer may construct some kind of a relational pattern out of the general probabilities of the case. The fundamental syntactic relations must be unambiguously expressed.—Sapir, 1921, p. 94

Speakers in every language talk about objects and actions and need to specify the relations among them. Word order is one device that allows them to do this (Greenberg, 1966; Hawkins, 1983). Although languages vary in how much they rely on surface order, all languages have some basic canonical arrangement (though see Mithun, 1992, for an alternative view). Moreover, languages may differ with regard to the range of permissible orders they exhibit. Some, like English, maintain a rigid SVO order, while others, like Warlpiri, a language of central Australia, allow all logically possible word order combinations (Hale, 1992). Yet within this variability, linguists have uncovered typological regularities that suggest the influence of universal principles of grammar (Comrie, 1981; Greenberg, 1963; Hawkins, 1983; Keenen, 1976). Such principles are believed to operate as powerful constraints on the orderings found within and across the languages of the world.

The study of word order universals and word order variability has long captivated linguists and psychologists alike (Bloomfield, 1933; Sapir, 1921; Wundt, 1900/1970). The present research attempts to shed new light on this age-old issue by exploring a number of factors that might lead individuals to rely on consistent word order. We construct an experimental situation that simulates various aspects of language-creation by instructing adults, naïve to sign language, to describe a series of brief scenes to another person, using only gesture and no speech. Previously, using this procedure (Goldin-Meadow, McNeill, & Singleton, 1996), we found that English-speaking adults not only invent discrete gestures to represent semantic elements in the scene, but also produce those gestures in strings characterized by consistent order. Moreover, the order the adults use is not the order found in typical English sentences—it is their own invention. Consistent ordering of semantic elements thus appears to be found, not only in languages handed down from generation to generation, but also in communication generated on the spot. In this paper, we use the gesture creation paradigm to explore factors that influence a communicator's reliance on word order. Before outlining our studies, we review the literature demonstrating the importance of order to language processors, language learners, and language creators.

1.1. The importance of word order in language processing, learning, and creation

Word order is one of the primary devices languages offer speakers to express who does what to whom. For example, the sentence "Oscar hits Felix" describes a very different situation from "Felix hits Oscar." Since languages vary in the orders they permit, the particular order of semantic elements that a language offers its speakers may or may not reflect the real-world sequencing of those elements. Within the constraints imposed by their language, speakers are more likely to produce surface structures that mirror the real-world sequencing of events (Clark & Chase, 1972; Osgood, 1983), and to process those sentences more readily than sentences whose order varies from real-world sequencing (Osgood & Bock, 1977; Sridhar, 1988).

Whether or not a particular order mirrors a natural sequence, there is evidence that placing a word in a particular position in a sentence can alter the psychological status of that word. For example, Gernsbacher and Hargreaves (1992) found a reliable advantage in the speed at which a name is recalled when the person is mentioned first, rather than last, in a sentence. Bock and Irwin (1980) found that words at the beginning of a sentence serve as better cues for recall than words in either the middle or end positions of a sentence. Finally, in studies that explore the neurological basis of syntax, a variety of measures indicate differences in how words are processed. For instance, placing content words at the front of a sentence appears to have measurable effects on ERP waveforms (Osterhout, 1994); placing words at the end and middle of a sentence appears to have identifiable effects on measures of eye fixation (Just & Carpenter, 1987). Taken together, these results provide empirical support for the view that the order in which words appear in a sentence has implications for the way those words are processed.

Word order also plays a central role in language acquisition. It is one of the first linguistic devices that children use to encode semantic roles (Bloom, 1970; Brown, 1973)—even if the language to which the child is exposed shows great flexibility in word order. For example, American Sign Language permits a variety of orders in addition to the unmarked SVO order. Despite this flexibility, deaf children learning ASL from their deaf parents use consistent order as a syntactic device for marking role early in development (Hoffmeister, 1978; Newport & Ashbrook, 1977; Newport & Meier, 1985). Order appears to be so pervasive in early language-learning that Slobin (1985, p. 1192) has suggested children come to the language-learning situation with a disposition to notice and store sequential orders of classes of elements.

Word order is a robust aspect of language-learning not only during childhood but later in life as well. Learners acquiring either a first (Newport, 1990) or second (Johnson & Newport, 1989) language beyond childhood display the same reliance on word order as early learners. Indeed, Newport

(1990) found that while age of first exposure to a language had a systematic effect on how well learners mastered *morphology* in that language, it had absolutely no effect on how well they learned *word order*—all learners were at ceiling on the word order tasks.

Not only is word order learned easily and effortlessly in the presence of a language model, it is also introduced into communication systems created de novo, over both long and short time-spans. As an example, over generations, word order is introduced during the process of expanding a simplified pidgin into a more complex language, a creole (Romaine, 1992). As a second example, on a time-span of months to years, word order is introduced when deaf children are forced by their circumstances to create their own gesture systems during childhood. Deaf children born to hearing parents are often not exposed to a conventional sign language until adolescence. Moreover, if their hearing losses are so profound as to preclude the acquisition of spoken language, they are unable to profit from the conventional spoken language that surrounds them. Despite their lack of access to a usable conventional language model, these deaf children invent gesture systems, which they use to communicate with the hearing individuals in their worlds (Feldman, Goldin-Meadow, & Gleitman, 1978; Goldin-Meadow & Mylander, 1984). Importantly from our point of view, the children use consistent order when producing gestures for particular semantic elements, and introduce this device into their communication systems whether they are growing up in American or Chinese cultures (Goldin-Meadow & Mylander, 1998). As a final example, on a time-span of minutes to hours, hearing adults impose order onto the spontaneous gestures they create when asked to describe a series of videotaped events using their hands and no words. Moreover, the order they use is not a carry-over from English, their spoken tongue, but appears to be invented at the moment (Goldin-Meadow et al., 1996).

Why is word order found in all of these varied communication situations? The language-creation examples make it clear that learners need not have a language model to introduce word order into their communication systems. The deaf children did so despite the fact that they had no access to conventional language. Moreover, the spontaneous gestures that the children's hearing parents produced when they talked to their deaf children did not display consistent order (Goldin-Meadow & Mylander, 1983), suggesting that the children did not even have an *unconventional* model for the order they introduced into their gestures. In addition, when asked to rely solely on gesture, the hearing English-speaking adults generated a non-English gesture order, suggesting that adults do not necessarily rely on the word order of their first language when generating a second "language." The fact that both adults and children exploit order when learning or creating a first or second language suggests that word order is not solely the product of a young, or of an old, mind.

Two other, not mutually exclusive, factors might play a role in encouraging communicators to rely on word order—communicative and cognitive factors. Word order may be essential in conveying information effectively to another person. If so, aspects of the communication situation itself might make it more, or less, likely that speakers will rely on order in their communications. Under this hypothesis, it is the relationship speakers and listeners hold to one another that makes ordering an essential strategy in communication. Alternatively, speakers may naturally parse events into a sequence of elements when apprehending those events. If so, order in communications may reflect this processing strategy. Under this hypothesis, it is qualities of the speaker that make ordering an inevitable component of communication.

Using the gesture creation paradigm developed by Goldin-Meadow et al. (1996), we ask in Study 1 whether consistent ordering is robust across a variety of communicative arrangements, arrangements that might be expected to shape how communication is structured (whether the communication partner is permitted to give feedback; whether the information to be communicated is present in the space that recipient and gesturer share; whether the gesturer assumes the role of gesture receiver as well as gesture producer). We will find that, not only is consistent ordering of semantic elements robust across the range of communication situations, but the same non-English order appears in all contexts. Study 2 explores whether this non-English order is bound to a communicative context at all; that is, whether it arises only when a person attempts to share information with another—a feature common to all of the conditions in Study 1—or appears in noncommunicative contexts as well. Finding the same non-English order in a non-communicative context would suggest that the order is not driven solely by communication in the manual modality, but may be a more general property of human thought.

2. Study 1

Human communication embraces a variety of definitions, from the transmission of information about objects and events in the world, to expressions of emotion, great literature, and poetry. Language serves the function of communication by adopting conventions specifying the range of forms and meanings possible to a particular social community. According to Clark (1996), face-to-face communication is the first, and most basic, setting of language use; it is immediate, reciprocal, and informative. Languages evolved, and continue to evolve, through this direct interplay between users. Yet within this moment-to-moment exchange, there exists a trade-off between the social need to be understood and the personal need to economize effort (Zipf, 1949).

The question we address here is how, in the absence of speech, do people impose organization on thoughts conveyed to another person. In Study 1, we hypothesize that a regular order will emerge in response to general pragmatic constraints that commonly operate in normal face-to-face interaction. Specifically, we test the hypothesis that participants will structure their gestures in a consistent and non-random fashion when communicating with another person who appears to be actively engaged in processing the message. We base this hypothesis on a large body of literature suggesting that language is fundamentally collective and collaborative in nature (Clark, 1992, 1996; Krauss & Fussell, 1991). Moreover, we assume that if the principal purpose of communication is to be understood by another person, then information concerning who-does-what-to-whom should be an essential ingredient of that process (though for a different view of the role of the listener, see Dell & Brown, 1991).

A number of studies provide insight into the nature of the collaborative communication enterprise. First, speakers are remarkably sensitive to the informational needs of their listeners. For instance, Clark and Wilkes-Gibbs (1986) found that speakers adjust the amount of detail they provide in establishing reference as a function of the common ground, or mutual knowledge, they share with their partner. This ability to take the perspective of the other person is found to play an important role in the comprehensibility of the message (Krauss, 1987). One prediction from these findings, then, is that ordering information in a sentence may be less essential to the comprehension of the listener when both speaker and listener witness the same event.

Other studies suggest that speakers also obtain and make valuable use of feedback from their listening partners. Clark and colleagues (Clark & Schaefer, 1987; Schober & Clark, 1989) found that listeners participate actively to indicate acceptance—by prompting the speaker for more information, expanding on what was previously said, and answering questions that the speaker poses. Moreover, speakers monitor their partners from moment to moment for evidence that they have been understood (Clark & Brennan, 1991). This tendency to continually update information about the listeners' state of understanding is likely to shape the speaker's ideas concerning the ordering of elements in a sentence when faced with a non-comprehending listener. A second prediction, then, is that in the absence of listener feedback, speakers might be less consistent in the ordering patterns they create.

Finally, speakers also find ways to promote coordinated understanding through their prior experience with a task. For example, Anderson and Boyle (1994) found that when speakers were given a second chance to guide an uninformed listener through a schematic map task, they appeared to change their assumptions about their partner's ability to interpret new information. We might expect that speakers would show even greater sensitivity to the experience of the listener if they themselves had actually been a listener in that situation; that is, if they have had the opportunity to assume the role of both producer *and* receiver. A third prediction, then, is that assuming both communicative roles will be a factor in the stability and perhaps the shape of the ordering patterns that speakers produce.

The present experiment examines the effect of these factors on the real-time construction order of elements in a message. We adopt an experimental paradigm based on previous research by Goldin-Meadow et al. (1996). In their studies, hearing adults with no sign language experience were asked to communicate information to another person using gesture in the absence of speech. Experiment 1 extends this design by manipulating the presence or absence of each of three factors: (1) The knowledge that speakers have in common with the communication partner, (2) the feedback that the partner provides to the speaker, and (3) the opportunity for the gesturer to serve as a receiver as well as a producer of messages. The central question addressed in this study is whether individuals will establish consistent ordering patterns to express relations between elements when communication is restricted to the manual modality.

3. Method

3.1. Participants

Sixty-four English-speaking students (half male and half female) from the University of Chicago were paid to participate in the experiment. All participants had normal hearing and no prior knowledge of ASL or other conventional sign language. Participants were randomly assigned to one of eight dyads across five different experimental conditions. Within each dyad, participants were assigned to the role of either gesturer or recipient. For two of the conditions, the experimenter or a confederate assumed the recipient role. In these conditions, then, only eight participants were recruited. In the remaining three conditions, 16 participants were recruited. Thus, there was a total of 40 gesturers and 24 recipients who participated in the experiment.

3.2. Materials and procedure

We adapted the Verbs of Motion Production (VMP) test (Supalla et al., in press), which was designed to assess knowledge of American Sign Language, to investigate the spontaneous gestures of hearing adults with no sign language experience. The VMP is an elicited production test composed of 120 very brief (1–2 s) scenes on videotape. Each scene is animated and entirely silent. In the present study, participants saw 40 of the 120 scenes. After each scene, one person assigned to the role of gesturer was asked to describe what happened using only gesture and no speech. Gesturers were told that they could use their hands, face, and body to communicate, but could not speak or make other sounds (e.g., "moo"). Nor were they allowed to manipulate props or get out of their seats. The partners who received the gestures also were instructed to refrain from talking or rising from their seats. No limit was placed on how long or how many times gesturers were allowed to describe the objects or events in each scene.

Half of the 40 animated scenes consisted of one object moving in varying paths and manners of motion across space (e.g., a porcupine moving along a zigzag path). The other half consisted of two objects: One stationary and one moving, and a crossing-space action (e.g., a doll jumping into a hoop). None of the scenes was related to any other, although the same object or action might appear more than once in a different scene (e.g., a hoop jumping over a tree). The present experiment focuses specifically on the 20 two-object vignettes that involve both a stationary and moving object.

3.3. Design

Individuals were randomly assigned to one of five conditions. As indicated in Table 1, the conditions vary according to the presence or absence of three separate factors: (1) Shared vs. New Knowledge, (2) Feedback, and (3) Switching Roles. In addition, although all conditions consisted of two people—one gesturer and one recipient—only in Conditions II, IV, and V were both participants naïve to the experiment. Specifically, in Condition I, the experimenter assumed the role of the recipient and in Condition III, a confederate assumed the role. These conditions are detailed further below.

3.3.1. Shared vs. new knowledge

In each condition, recipients were either familiar or unfamiliar with the message that the gesturer was about to communicate. In Conditions I and II, the gesturer and recipient both watched the videotape of the vignette; thus, both partners knew the gist of the message to be communicated and no new knowledge was imparted (−NK). In this sense, the recipient did not receive new information from the gesturer. Note that this condition is reminiscent of here-and-now conversations in which one partner comments on a scene that both have witnessed. In contrast, in Conditions III–V, only

Table 1
Design of the five experimental conditions in Study 1

Condition	New knowledge	Feedback	Role switch
I	−NK	−F	−RS
II	−NK	+F	−RS
III	+NK	−F	−RS
IV	+NK	+F	−RS
V	+NK	+F	+RS

the gesturer watched the videotape; thus, the recipient and gesturer did not share background information on the scene. The information that the gesture conveyed was consequently new to the recipient (+NK).

3.3.2. Feedback

In addition to whether or not knowledge of the scene was shared with the recipient, the five conditions also varied according to whether or not the recipient was instructed to supply feedback to the gesturer. Here, feedback refers strictly to nonverbal responses, since speech was not permitted in the experiment. Thus feedback includes all facial, manual, or vocal (grunts, sighs, etc.) information that might cue the gesturer as to whether the message was comprehended by the receiver. Recipients were told that they could use their face or hands to ask questions or show understanding. No restrictions were set on how often, or when, this feedback might be provided. Typically, an exchange ended when the recipient indicated his or her understanding of the message with a nod or 'okay' sign.

As shown in Table 1, recipients in Conditions I and III provided no feedback to the gesturer (−F); recipients in Conditions II, IV, and V did provide feedback (+F). Importantly, in the no feedback conditions, we used the experimenter in Condition I and a confederate in Condition III to act as the communicative partner. This enabled us to maintain control over possible inadvertent "leaks" of information by "real" participants, that is, those naïve to the experiment. Confederates were trained to avoid expressions that would convey understanding of the vignettes, such as a smile, a nod, or a furrowed eyebrow. Their instructions were to look attentive and to maintain a neutral face. Gesturers, in turn, were instructed to continue their efforts to convey the information to their partner until they felt their partner understood.

3.3.3. Switching roles

Finally, the conditions differed in whether the participants took a turn as both the receiver *and* producer of gesture. Specifically, in Condition V, partners were randomly assigned to start the experiment as either the gesturer or the recipient. Partners then switched roles after each block of 10 vignettes (+RS). Thus there was a total of four blocks, two for each participant as the gesturer and two as the recipient. In the remaining conditions, where partners did not exchange roles (−RS), one participant was always the gesturer and the second was always the recipient.

3.3.4. Mutual understanding

To assess comprehension of information received by the communication partner, both the gesturer and the recipient in each dyad were instructed to write down what they thought the gesturer had communicated immediately following each vignette. Participants were encouraged to use full sentences rather than to list isolated words. The one exception to this procedure was Condition I, in which the experimenter assumed the role of recipient. In that condition, only the gesturer was asked to record what he or she had attempted to communicate.

3.4. Coding

All sessions were videotaped and later analyzed, first, for the number of gestures produced (i.e., tokens) and, second, for the production of gesture combinations, or "strings." In addition, each gesture token was coded for both its form and meaning, as described below. Participants generated many different kinds of gestures in their efforts to describe the scene to their partner. For example, the following gestures were used by one participant to describe a scene in which a fireplace appeared as the stationary object: (1) Uses finger to trace the rectangular shape of a fireplace in the air; (2) pantomimes striking match on matchbox; (3) waves arms up and down to indicate flames shooting up; (4) pantomimes warming hands in front of fire; (5) places one arm horizontally in space to indicate mantle of fireplace. All five of these gestures were counted as referring to a single stationary object.

Two non-signing coders were given extensive training on how to identify and interpret individual gestures in the context of the videotaped scenes. As a result, they had little difficulty determining what the scenes were or which gestures referred to which components. When ambiguity did arise, coding decisions were discussed and mutually agreed upon.

3.4.1. Gesture form

The form of individual gestures was determined by the shape and orientation of the hand and by the trajectory of motion. For example, to describe an airplane flying around a tree, one participant placed his right hand, palm facing down in the horizontal plane, to represent the airplane and his left hand, palm facing outward in the vertical plane, to represent the tree. Holding these handshapes, he then moved his right hand in a semi-circular path around his left hand.

3.4.2. Gesture meaning

The meaning of individual gestures was determined in the context of the three elements that appeared in each scene: The moving object, the stationary object, and the action. If the participants produced a gesture that captured some aspect of an element in the scene, it was coded as representing that element. For example, one participant sucked her thumb to represent a baby, another pantomimed cradling an infant in her arms.

Participants also represented information about an element by holding a handshape, which captured some characteristic of that element in the air, and then incorporating that handshape into the action gesture. For example, to describe the motion of a baby waddling across the screen, one participant held a leg-shaped V-hand in the air and then used that handshape in the rocking and moving action that came next. As another example, one that involves gestures for both the stationary object and the moving object, one participant described a wreath falling from above a fireplace by first placing a left flat palm (representing the mantel of the fireplace) in the air, then holding up a right handed circle (representing the wreath) over her left palm, and finally moving the circle downward past the palm to represent the fall.

3.4.3. Gesture strings

Individual gestures were frequently combined with one another to produce "strings"—sentence-like constructions in which gesture order could be used to convey semantic relations between elements. Two or more gestures were considered a string if they were connected by a continuous flow of movement. The end of a string was coded if the hand retracted or relaxed, or if there was a pause longer than a second between two distinct gestures.

3.4.4. Comprehension measure

To assess the accuracy with which the recipient understood the message conveyed by the gesturer, we determined the percent agreement for each dyad for the three elements (stationary object, moving object, and action) in all two-object scenes. Agreement was broadly defined to include synonyms (e.g., girl and female) as well as membership in the same superordinate category (e.g., food and egg). Imprecise responses, such as "object" and "thing," were considered ambiguous and were not counted as agreements. Comprehension data were analyzed only in Conditions II, IV, and V, where both participants were naïve to the experiment.

3.4.5. Reliability

For each of the five conditions, two coders established reliability on half the participants for five out of the 20 two-object scenes. Thus, reliability consisted of a subset of 100 scenes. There was 82% agreement for classifying a gesture as an action, a moving object, or a stationary object. Using Cohen's κ to correct for chance agreement, interrater agreement was .76. Percent agreement for determining the number and ordering of elements in a string was .88 (Cohen's $\kappa = .87$).

Reliability for the comprehension data was based on half of the two-object scenes for two randomly selected participants in each of the three conditions for which we analyzed agreement between the gesturer and recipient (Conditions II, IV, and V). Reliability was 91% for the moving object, 88% for the stationary object, and 93% for the action gesture (Cohen's $\kappa = .89$, .83, and .72, respectively). The main source of discrepancy between reliability coders arose in deciding whether participants' written responses were semantically equivalent, for example, ring and hole.

4. Results

4.1. Comprehension of gestures

In Experiment 1, participants assigned to the role of gesturer were responsible for making themselves understood while gesturing to their communication partner about the scenes they had witnessed. For the gestures to work successfully as a communication system, it was necessary both to generate a lexicon of distinctive gestures and to establish a systematic way of indicating the relations between them. Our measure of understanding was based on the proportion of written responses provided by the recipient that matched the written responses of the gesturer, following the gesturer's attempt to describe each vignette without speech. Table 2 presents the mean percent agreement (and standard deviation) between partners for the three conditions in which a naïve recipient was used (II, IV, and V). In the other two conditions, the experimenter or a confederate was the recipient and these conditions were excluded from this analysis. The data are presented separately for the Moving object (M), the Stationary object (S), and the Action (A).

The results indicate that recipients were fairly accurate in interpreting the meaning of their partners' gestures. Across conditions, the mean percent agreement for gestures representing the Moving object, Stationary object, and Action was .59, .60, and .74, respectively. We conducted an analysis of variance with one between-subject factor (Condition) and one within-subject factor (Element). The results indicated a significant main effect of both condition, $F(2,20) = 10.1$, $p < .001$, and element, $F(2,21) = 5.86$, $p = .01$, but no interaction. In particular, post hoc comparisons using Tukeys hsd ($\alpha = .05$) show that agreement between partners was reliably greater for the Action gestures than for either the Moving or Stationary objects, thus suggesting a somewhat greater transparency of verb-like rather than noun-like gestures. In addition, as expected, mutual understanding was significantly improved in Condition V where partners switched roles (+RS) to

Table 2
Mean percent agreement between gesturer and recipient on gestures generated to represent Moving objects, Stationary objects, and Actions[a]

Condition[b]	Moving object	Stationary object	Action
II (−NK, +F, −RS)	.61	.68	.80
	(.27)	(.27)	(.08)
IV (+NK, +F, −RS)	.45	.42	.61
	(.16)	(.19)	(.22)
V (+NK, +F, +RS)	.71	.70	.82
	(.09)	(.07)	(.06)

[a] Standard deviations are below the means in parentheses.
[b] Data are presented for the three conditions in which a naïve recipient participated.

become both the producer and receiver of gesture, relative to Condition IV where there was no opportunity to switch roles (−RS). Moreover, when feedback was made available, mutual understanding was better in Condition II, where knowledge was shared between partners (−NK, +F) than in Condition IV, where new knowledge was introduced (+NK, +F).

4.2. Number of gestures

The gesturers referred to the objects and actions in the scenes they witnessed by capturing an aspect of those objects or actions in their gestures. As indicated earlier, a gesturer could produce more than one gesture to refer to the same object or action, thus increasing the chances that the recipient would arrive at the correct referent. We predicted that gesturers would tend to use a large number of gestures when the recipients had no prior knowledge of the information to be conveyed (i.e., when the information was all new, +NK), and when the recipients could signal their lack of comprehension (i.e., when they could provide feedback, +F). Taking a turn as recipient (+RS) might also make it more likely that the gesturers would produce a large number of gestures simply because they knew precisely what was involved in interpreting gesture.

The extent to which individual gesturers in each of the five conditions attempted to produce gestures for the objects and actions in the scenes is reflected in the total number of gestures they produced (i.e., tokens; tokens included all gestures that a participant used to refer to a particular object, whether those gestures captured different aspects of the referent or repeated the same aspect). In the analyses that follow, we evaluate how each of the three factors we varied (introduction of new knowledge, feedback, and role switching) influenced the frequency of gesture tokens. We first compare the gesturers' responses in Conditions I–IV to determine the impact of knowledge and feedback, using an analysis of variance with two between-subjects factors, knowledge (+NK, −NK) and feedback (+F, −F). We then conduct an independent samples t test to examine, separately, the effect of role switching in Condition V (+NK, +F, +RS) relative to Condition IV which differed from V only in role switch (+NK, +F, −RS).

The data are presented in the left hand column of Table 3 as the mean number of gesture tokens per scene for each of the eight gesturers per group. Results of the analysis of variance revealed effects of new knowledge, $F(1,28) = 44.5$, $p < .0001$, and feedback, $F(1,28) = 21.1$, $p = .001$, and no interaction between knowledge and feedback. Participants produced more gesture tokens in Conditions III and IV where new information was introduced to the recipient (+NK) than in Conditions I and II where the information was previously known to the recipient (−NK). Participants also produced more gestures in Conditions II and IV where feedback was given (+F) than in Conditions I and III where it was not (−F). Finally, an

Table 3
Mean number of gesture tokens and gesture strings per vignette as a function of condition[a]

Condition	Gesture tokens	Gesture strings
I (−NK, −F, −RS)	2.2	.81
	(.55)	(.17)
II (−NK, +F, −RS)	4.2	1.14
	(1.82)	(.97)
III (+NK, −F, −RS)	5.6	1.17
	(2.19)	(.29)
IV (+NK, +F, −RS)	9.5	2.19
	(2.26)	(.45)
V (+NK, +F, +RS)	10.9	2.50
	(3.89)	(.72)

[a] Standard deviations are below the means in parentheses.

independent t test assessing the effect of switching roles on gesture tokens in Conditions IV vs. V found no significant difference, $t(14) = .89$. That is, contrary to our prediction, taking turns as gesturer and recipient (+RS) did not effectively alter the number of gestures produced by the participants. Thus, two of the communicative factors we varied—how much information was new to the recipient, and whether the recipient gave feedback—had effects on the number of gestures participants produced. The third factor—whether the gesturer both produced and received the gestures—did not.

4.3. Gesture strings

Table 3 also presents the mean number of gesture strings produced by the eight gesturers per group for each scene. The gesture strings were used to express the semantic relations between elements in a scene and, in this sense, resembled sentences. A single gesture string was defined as two or more gestures connected by a continuous flow of information. Again, we used an analysis of variance with two between-subjects factors (new knowledge and feedback) to compare Conditions I–IV. We found main effects of knowledge, $F(1,28) = 40.7$, $p < .001$, and feedback, $F(1,28) = 37.8$, $p < .001$. Participants produced a larger number of gesture strings when new information was introduced to the recipient (Conditions III and IV, +NK) and when the recipient provided feedback (Conditions II and IV, +F). Moreover, the two communication factors had an interactive effect, $F(1,28) = 10.0$, $p = .007$. Gesturers produced reliably more strings in Condition IV where both communicative factors were in play (+NK, +F) than in any of the other three conditions ($p < .01$, Tukeys hsd). Finally, an independent t-test assessing the effect of switching roles on gesture strings in Conditions V vs. IV found no significant difference, $t(14) = .99$, *ns*. Thus, sentence-like constructions occurred most frequently in face-to-face interactions that combined imparting new knowledge with the use of feedback

(+NK, +F). The production of strings was not influenced, however, by having the opportunity to assume the role of both gesturer and recipient (+RS).

4.4. Ordering the elements in a gesture string

We next turn to the central question of the study—does communicative context influence whether semantic elements are gestured in a particular order and, if so, does it influence the particular order in which they appear? Although each scene contained three different element types—the Moving object, the Stationary object, and the Action—gesturers could produce fewer than three gestures within a single string, or more than three. That is, gesturers might sometimes omit one of the three semantic elements or choose to produce an element more than once. In order to reduce the complexity of analysis that would result from allowing an unrestricted number of elements in a string, we focused the remaining analyses on 2- and 3-elements strings that included an Action gesture as one of the elements in the sentence. This allowed for four possible orderings among the 2-element strings (MA, SA, AM, and AS) and six possible orderings among the 3-element strings (SMA, MAS, SAM, MSA, AMS, and ASM). Table 4 displays the total number of 2- and 3-element gesture strings and the proportion of all gesture strings that comprise the database for our ordering analyses (i.e., the analyses in Figs. 1 and 2).

4.4.1. Two-element gesture strings

Figs. 1a–e present the proportion of gesture strings of each ordering type (MA, SA, AM, and AS) that the participants in each of the five conditions produced. Several points are apparent. First, across all conditions, there is an extremely low proportion of 2-element gesture strings that begin with

Table 4
Total number of 2- and 3-element strings that include an Action gesture and the proportion of total strings used in the ordering analyses as a function of condition[a]

Condition	2-Element	3-Element	Total strings (%)
I (−NK, −F, −RS)	66 (3.98)	53 (4.45)	.95
II (−NK, +F, −RS)	46 (3.89)	71 (3.91)	.67
III (+NK, −F, −RS)	24 (2.12)	58 (4.17)	.46
IV (+NK, +F, −RS)	54 (4.73)	116 (4.33)	.51
V (+NK, +F, +RS)	111 (6.82)	70 (3.56)	.48
Total	301	368	

[a] Standard deviations are below the means in parentheses.

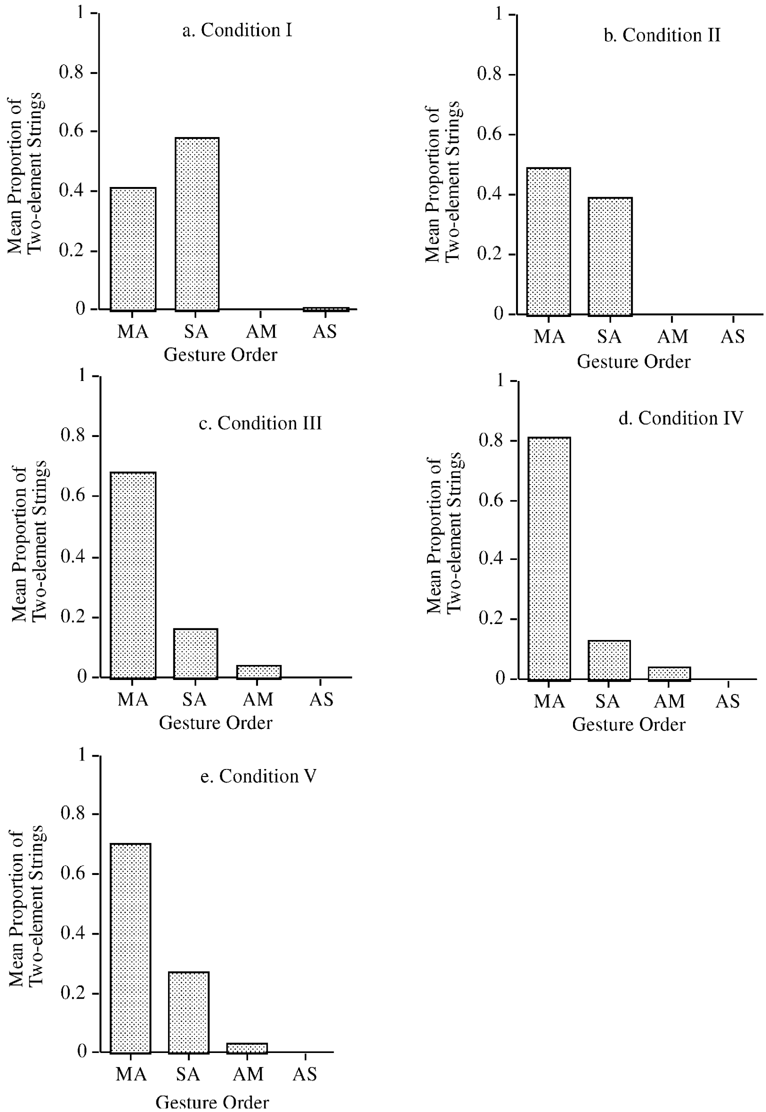

Fig. 1. Mean proportion of 2-element strings that participants produced with an MA, SA, AM, or AS ordering in: (a) Condition I (−NK, −F, −RS), (b) Condition II (−NK, +F, −RS), (c) Condition III (+NK, −F, −RS), (d) Condition IV (+NK, +F, −RS), and (e) Condition V (+NK, +F, +RS). M, Moving object; A, Action; S, Stationary object.

an action (less than .05 in any condition for AM or AS combined). In virtually all cases when gestures for Moving objects or Stationary objects are produced in a gesture string with gestures for Actions, the Moving or Stationary object precedes the Action (MA or SA). This ordering is not surpris-

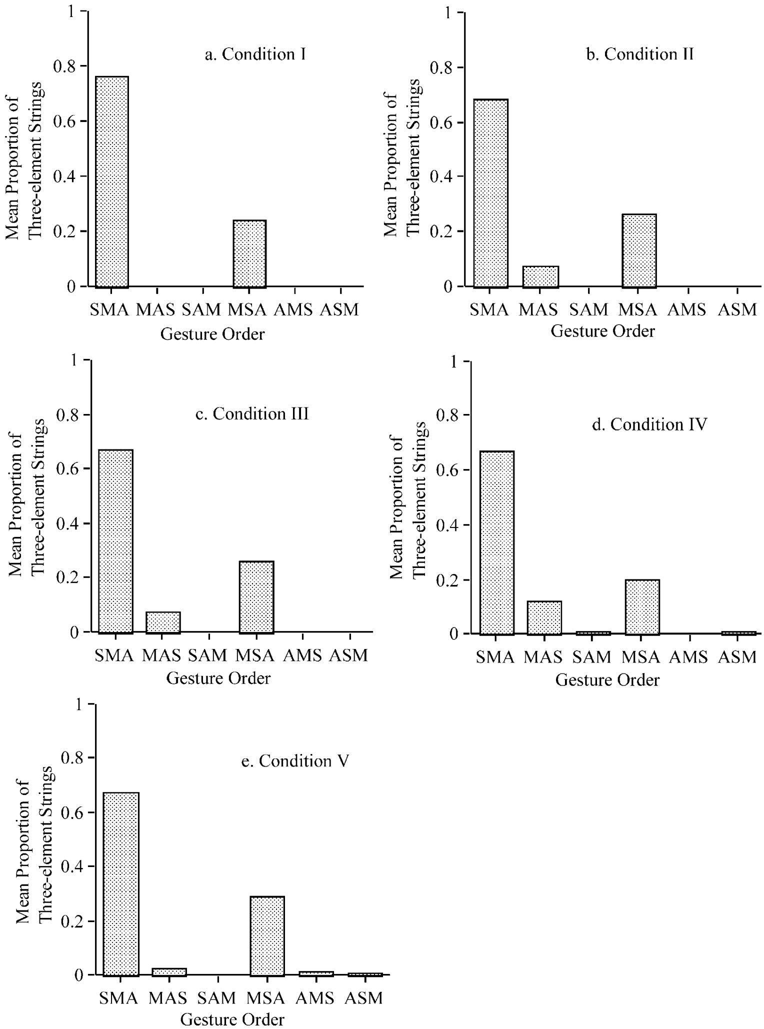

Fig. 2. Mean proportion of 3-element strings that participants produced with an SMA, MAS, SAM, MSA, AMS, or ASM ordering in: (a) Condition I (−NK, −F, −RS), (b) Condition II (−NK, +F, −RS), (c) Condition III (+NK, −F, −RS), (d) Condition IV (+NK, +F, −RS), and (e) Condition V (+NK, +F, +RS). M, Moving object; A, Action; S, Stationary object.

ing for the Moving object (girl-jump), as it mirrors the canonical ordering of elements in English ("the girl jumps"). However, the ordering of the Stationary object before the Action (hoop-jump) is not typical of English—the common pattern in English places the Stationary object after the Action ("jumps into hoop"). All eight of the participants in each of the five conditions produced more MA strings than AM strings, and more SA strings than AS strings. Thus, the communicative factors we varied—imparting new knowledge, receiving feedback, switching roles—had no effect on these particular gesture orders produced by the participants in their two-gesture strings.

However, one of the communicative factors—imparting new knowledge—did affect how often the participants produced MA and SA strings. We conducted separate analyses of variance for MA and for SA, and found a main effect of new knowledge for both ($F(1,28) = 6.10$, $p = .02$ for MA, and $F(1,28) = 9.92$, $p = .004$ for SA). Participants who were imparting new knowledge to their recipients (Conditions III and IV) produced more MA strings and fewer SA strings than participants who were not (Conditions I and II). Feedback had no effect, and switching roles had an effect only on the Stationary object (SA strings were more frequent when the gesturers and recipients switched roles than when they did not, $t(14) = -2.19$, $p = .05$). In sum, the communicative factors we varied had an effect on *whether* the participants produced strings containing the Moving object or the Stationary object. However, these factors had no effect on *where* the Moving and Stationary objects were produced in relation to the Action—both were consistently produced before the Action in all five conditions.

4.4.2. Three-element gesture strings

Figs. 2a–e present the ordering data for the three-element gesture strings produced by the participants under each of the five experimental conditions. The pattern of results is virtually identical, independent of condition. In all five of the communicative contexts, participants much preferred to place gestures for the Stationary object in initial position, gestures for the Moving object in medial position, and gestures for the Action in final position (SMA). All eight of the participants in Condition IV used this SMA order more than all of the other possible orders combined, as did seven of the eight participants in Conditions II and V, and six of the eight participants in Conditions II and III. We conducted an analysis of variance with two between-subjects factors (new knowledge, feedback) to compare the use of SMA in Conditions I–IV, and found no significant effects ($F(1,28) = .33$, *ns* for new knowledge; $F(1,28) = .24$, *ns* for feedback). In addition, we conducted a t-test to explore the impact of role switching on SMA in Conditions IV vs. V, and again found no significant effects ($t(14) = -.04$, *ns*).

Figs. 2a–e also show that, like the 2-element strings, almost no gesture sentences began with an action gesture (AMS or ASM). Finally, and perhaps most surprising, in none of the conditions was English word order (MAS = "the girl jumps into the hoop") the most frequent gesture order—indeed only a very small proportion of the responses followed this canonical English order in any condition. Thus, as in the two-element gesture strings, none of the communicative factors we varied had any impact on *where* gestures for Moving and Stationary objects were produced in relation to gestures for Actions—the SMA ordering was robust across all five conditions.

5. Conclusion

In Study 1, we found that *quantity* of communication (both number of gestures and number of gesture sentences) was affected by the communicative factors we varied—whether new knowledge was imparted to the recipient, whether feedback was received from the recipient, and whether the same person took turns being gesturer and recipient. Contrary to our expectations, however, the communicative factors we varied had *no* effect on the form those gesture communications took—in particular, on the order in which gestures appeared in a string. Participants relied on the same SMA gesture order across all five conditions. Moreover the SMA gesture order the participants used did *not* mirror English word order. Thus, the participants did not simply translate English into gesture, but rather created their structured gesture sentences de novo. These consistent findings across experimental conditions suggest that the SMA ordering pattern is not shaped by the particular communication pressures that we manipulated.

What, then, does shape this ordering pattern? Interestingly, although the basic structure of ASL, like English, is subject-verb-object (Fischer, 1975; Liddell, 1980), SMA is an order found routinely when signers describe crossing-space events involving stationary objects (Supalla, 1982). SMA was also found in a gesture system developed by a woman who was the only deaf person in a hearing Amerindian reservation (Yau, 1985). Thus, the SMA order appears to be one that is naturally exploited in manual communication systems—in conventional sign languages handed down from generation to generation, as well as in spontaneous gesture systems created by individuals.

Our findings, taken together with observations of more conventional sign languages, suggest that the SMA order may be a product of communicating in the manual modality. If so, the order ought to disappear if we alter our experimental task so that it no longer makes use of the manual modality, and so that it no longer involves communication. Study 2 tests this hypothesis by exploring ordering in a non-communicative task that does not involve gesture.

6. Study 2

Freyd (1983) has suggested that certain knowledge structures emerge because of their social utility at the group level rather than at the individual level. For example, color is perceived along a continuous dimension, yet is made accessible to the minds of others by fitting individual experiences into discrete, conventional categories. Freyd might predict that the ordering regularities we observed in Study 1 emerged because the task was to impart information to another person. An alternative hypothesis, one that we examine in Study 2, is that the ordering relations found in communication are an outgrowth of the way we conceptualize and parse scenes, rather than an outgrowth of the communicative situation itself. We explore this possibility by removing the ordering task from the communicative realm.

The challenge in Study 2 was finding a task that involved symbols but did not involve communication. We met the challenge by making two changes to Study 1. First, we eliminated the presence of a social partner so that only one person was involved in the task. Second, we eliminated the use of speech or gesture as the means for conveying information about order. Instead, we introduced the use of pictorial symbols as the method for representing the elements in a scene.

Participants in Study 2 were shown the same scenes used in Study 1. After each scene, the participant was given three line drawings, each representing one of the three elements in the scene: The Moving object, the Stationary object, or the Action (the action was portrayed by directional lines as in cartoon drawings). Each picture was drawn on a separate transparency. Individuals were asked to "reconstruct the scene" by stacking the transparencies on top of one another. To discourage participants from sequencing the pictures in space (that is, from left to right), we instructed them to stack the pictures on an overhead projector, aligning the dots on the left-hand corner of each page. An example of the pictures is presented in Fig. 3.

We introduced one other manipulation into Study 2. We varied the location of the Moving object on the transparency: In some vignettes, the Moving object appeared in its initial state, prior to the action taking place; in others,

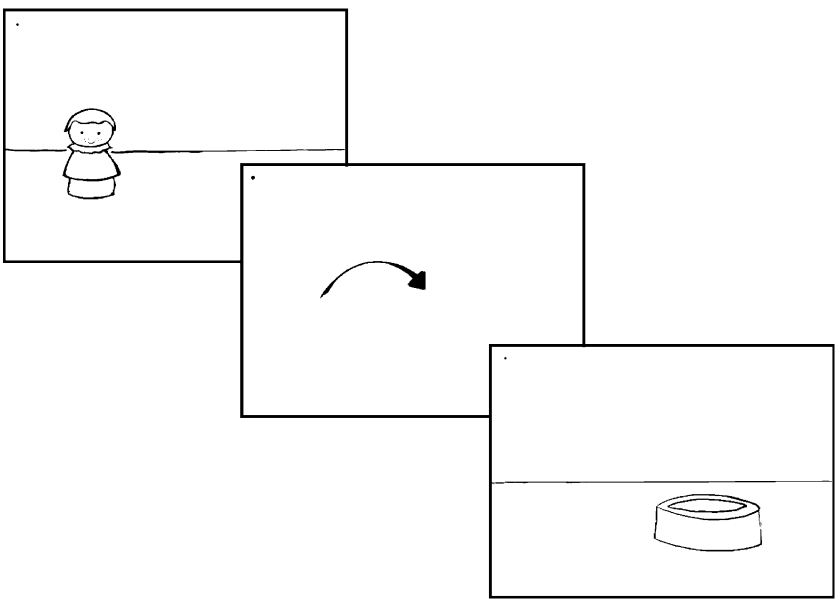

Fig. 3. An example of the three pictures used in Study 2. The girl represents the Moving object, the arrow represents the Action, and the hoop represents the Stationary object.

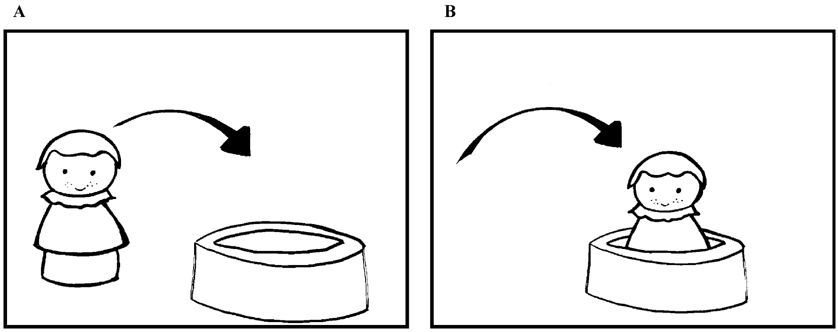

Fig. 4. The Moving object pictured: (A) in its Initial state before the action occurs and (B) in its Final state after the action has occurred.

the Moving object appeared in its final state, after the action had taken place. Fig. 4 presents an example. The panels show what the three transparencies would look like after they had been stacked. In Panel A, the doll figure, which represents the Moving object, is located outside the hoop in the position it occupied *before* the jumping action begins. In Panel B, the doll figure is shown inside the hoop in the position it occupied *after* the jumping action has been completed. If the semantic role that the Moving object plays is all that is important, asking the participant to reconstruct the scene in either its initial state or final state should have no effect on whether a consistent order is used, or what that order is. However, perspective on the event—that is, whether the participant considers the event before it occurs or after—might well influence how the scene is conceptualized. If so, then reconstructing the scene in either its initial or final state could affect whether a consistent order is used, and what that order is. To test these alternative hypotheses, we made the state of the Moving object a within-subject variable.

We examined how participants reconstructed the scene under three conditions. In the Self condition, participants were instructed to reconstruct the scene for themselves. This condition was designed to remove all components of the communicative task. For comparative purposes, we included two other conditions. Participants were instructed to reconstruct the scene for another person (the Other condition), or while talking to the experimenter (the Talking condition).

7. Method

7.1. Participants

Sixty undergraduates from Carnegie Mellon University volunteered to participate for research credit. All participants were native speakers of English. None was a participant in Study 1.

7.1.1. Materials and procedure

The set of 20 two-object scenes used in Study 1 was presented individually to participants. Participants were also presented with 20 sets of overhead transparencies, each containing three separate pictures, and an overhead projector. These 20 sets of transparencies corresponded to the 20 scenes in the vignettes. After viewing each scene on videotape, participants were asked to "reconstruct the scene" using the three transparencies representing the Moving object, the Stationary object, and the Action. The three transparencies were placed in a triangular arrangement in front of the participant, with placement counterbalanced for each person across set. Participants were instructed to place the transparencies, one by one, on the overhead projector and to align the dots in the upper-left hand corner. This required participants to stack the pictures rather than lay them out sequentially.

In half the pictures, the Moving object was pictured in its initial state—that is, before the action took place. In the other half, the Moving object was pictured in its final state—after the action had taken place. The state of the Moving object was rotated in blocks of five, such that all participants received first five sets of transparencies in one state and then five sets in the other state. The order in which the first block was presented was counterbalanced across participants.

7.2. Design

The participants were randomly assigned to one of three conditions. In the Self condition, participants were instructed simply to "reconstruct the scene" after viewing the videotape. The experimenter strategically sat behind a large screen and made it clear that she could not see what the participant was doing. To emphasize the point, participants were required to tell the experimenter when they were ready for the next set of pictures. This condition was intended to be non-communicative. In the Other condition, participants were instructed to "reconstruct the scene for someone else." Participants were told that another person, naïve to the experiment, would be viewing the videotaped session in another part of the study. The task of these naïve participants would be to describe what happened in the scene, based on the pictures the current participant placed on the overhead projector. No screen was used to block the experimenter's view. This second condition was intended to maximize the sense that the task was communicative in nature and thus contrasts with the Self condition. The third Talking condition was similar to the second in that there was no screen to block the experimenter's view. In addition, participants were instructed to describe what they were doing as they placed the overheads on the projector. The purpose of this condition was to determine the effect of English word order on the sequencing of pictures in an otherwise nonverbal task.

To assess whether our manipulation succeeded in convincing participants in the Self condition that the task was non-communicative, we queried participants in all three conditions during a debriefing period at the end of the experiment. We asked, "Did you feel like you were communicating to anyone else during this experiment?" Our concern was that despite our efforts to limit the interactive nature of the task, participants might respond to the videotaping in the session as if they were communicating to another person.

7.3. Coding

Participants' responses during the picture-arranging task were coded according to the order in which each set of three transparencies was placed on the overhead projector. No limit was imposed on how many times the stacks could be constructed, though few participants chose to rearrange the pictures once placed. Coding decisions were based on the final arrangement of pictures on the overhead; the picture placed on the bottom of the stack was considered the first of the series since it was most often the first to be picked up from the table and placed on the overhead.

In addition, we calculated the number of times participants in each condition responded "yes" to the manipulation check presented during the debriefing period.

7.3.1. Reliability

Two coders established reliability for half the participants randomly selected in each of the three conditions. Assessment was based on participants' ordering of the semantic elements in each set of 20 scenes. Intercoder agreement was 94% for the Self condition (Cohen's $\kappa = .88$), 93% for the Other condition (Cohen's $\kappa = .87$), and 96% for the Talking condition (Cohen's $\kappa = .96$).

8. Results

8.1. Manipulation check

Did the participants feel that they were communicating to someone when reconstructing the scene? The mean proportion of individuals answering "yes" to this question was .15 in the Self condition, 1.00 in the Other condition, and. 82 in the Talking condition. Thus, our manipulation was successful. Participants who were instructed to reconstruct the scene for themselves did not perceive the picture arranging task to be communicative in nature, whereas those instructed to reconstruct the scene for someone else, or describe what they were doing as they placed the pictures on the overhead projector, did.

8.1.1. Did the participants use a consistent order to stack the pictures?

We next ask whether the participants used a predominant ordering when arranging the pictures for themselves, for another, or while talking. To determine whether the individual participants in each condition used a predominant ordering, we first identified for each participant the ordering that he or she used most often. We then calculated the proportion of reconstructions in which the participant used that ordering. We did these calculations separately for sets in which the Moving object was pictured in its Initial state and for sets in which the Moving object was pictured in its Final state.

For pictures in which the Moving object was in its Initial state, participants in each condition used their predominant ordering very frequently in their reconstructions: 87% in the Self condition, 86% in the Other condition, and 89% in the Talking condition. Similarly, when the Moving object was in its Final state, participants in the Self, Other, and Talking conditions used their predominant ordering in 83, 89, and 87% of their reconstructions, respectively. Thus, condition did not affect how likely participants were to use a consistent ordering when stacking the pictures.

Surprisingly, condition did affect *which* predominant order the participants used. Table 5 presents the proportion of participants in each condition ($N = 20$ per condition) who used SMA, SAM, MAS, or MSA as their predominant stacking order (none of the participants used the remaining two orderings, ASM or AMS, predominantly). A very small number of participants could not be assigned a predominant ordering—this includes one participant when stacking pictures with the Moving object in Initial state, and three when stacking pictures with the Moving object in Final state. The data for these participants in these conditions are not included in Table 5; thus, the proportions in the table do not sum to 1.00 for all of the conditions.

We examined first the pictures with the Moving object in Initial state. The most striking result is that 80% of the participants in the Self condition stacked the pictures using the SMA ordering—precisely the ordering that

Table 5
Proportion of participants using a predominant order when the Moving object was pictured in its Initial or Final state as a function of condition[a]

Condition	Initial state				Final state			
	SMA	SAM	MAS	MSA	SMA	SAM	MAS	MSA
Self	.80	.05	.05	.05	.50	.40	.05	0
Other	.70	.05	.25	0	.20	.50	.20	0
Talking	.45	.05	.50	0	.25	.30	.45	0

[a] There were 20 participants in each condition. One participant in the Self condition used the SMA and SAM order equally often when the Moving object was in Initial state. One additional participant in the Self condition and two in the Other condition used SMA and SAM equally often when the Moving object was in Final state. These participants could not be assigned a predominant order. As a result, the proportions for these conditions do not sum to 1.00.

we found in the gesture sentences created by the adults in Study 1. In contrast, only 45% of the participants in the Talking condition used the SMA ordering. Not surprisingly since the participants in this condition were asked to talk while stacking their pictures, 50% of these participants followed the MAS ordering which mirrors canonical English word order ("girl jumps into hoop"). The number of participants who adhered to an SMA ordering rather than an MAS ordering differed significantly comparing the Self vs. Talking conditions (16 SMA and 1 MAS for Self vs. 9 SMA and 10 MAS for Talking; $X^2 = 7.16$, $df = 1$, $p = .007$). Participants in the Other condition fell in between along this dimension.

Turning next to pictures with the Moving object in Final state, we see that the major effect of varying the Moving object's state from initial to final was, in all three conditions, to increase the proportion of participants who used an SAM ordering. Across all three conditions, only 3 of the 60 participants used SAM when the Moving objects was in Initial state vs. 24 of 60 when the Moving object was in Final state ($X^2 = 19.12$, $df = 1$, $p < .001$). Interestingly, however, the difference between the Self and Talking conditions in the other two orders remained evident in the Final state pictures: The number of participants who adhered to an SMA ordering rather than an MAS ordering still differed significantly comparing the Self vs. Talking conditions (10 SMA and 1 MAS for Self vs. 5 SMA and 9 MAS for Talking; $X^2 = 5.69$, $df = 1$, $p = .02$).

In addition to analyzing the data in terms of predominant ordering patterns, we also examined all of the responses that the participants in each of the three condition produced. Fig. 5 presents the proportion of vignettes in which participants stacked pictures according to one of the six possible orderings when the Moving object was in its Initial state (a–c, graphs on the left), and when the Moving object was in its Final state (d–f, graphs on the right). In each set of graphs, the top panel displays the data for the (a,d) Self condition, the middle panel displays the data for the (b,e) Other condition, and the bottom panel displays the data for the (c,f) Talking condition.

The patterns we saw in the individual analyses are apparent in the group analyses as well. We conducted an analysis of variance with one between-subjects factor, Condition (Self, Other, Talking), and one within-subjects factor, State of the Moving Object (Initial, Final), on each of the three predominant ordering patterns that the participants used (SMA, MAS, and SAM).

Beginning with SMA, we found a marginal effect of Condition, $F(2, 57) = 2.91$, $p = .06$, a robust effect of State, $F(1, 57) = 19.83$, $p < .0001$, and no interactions. Participants used SMA significantly more often when the Moving object was in its Initial state than when it was in its Final state. In addition, planned comparisons revealed that participants in the Self condition used SMA significantly more often than participants in the Talking condition ($F(1, 57) = 5.61$, $p = .02$), just as we found in the individual analysis.

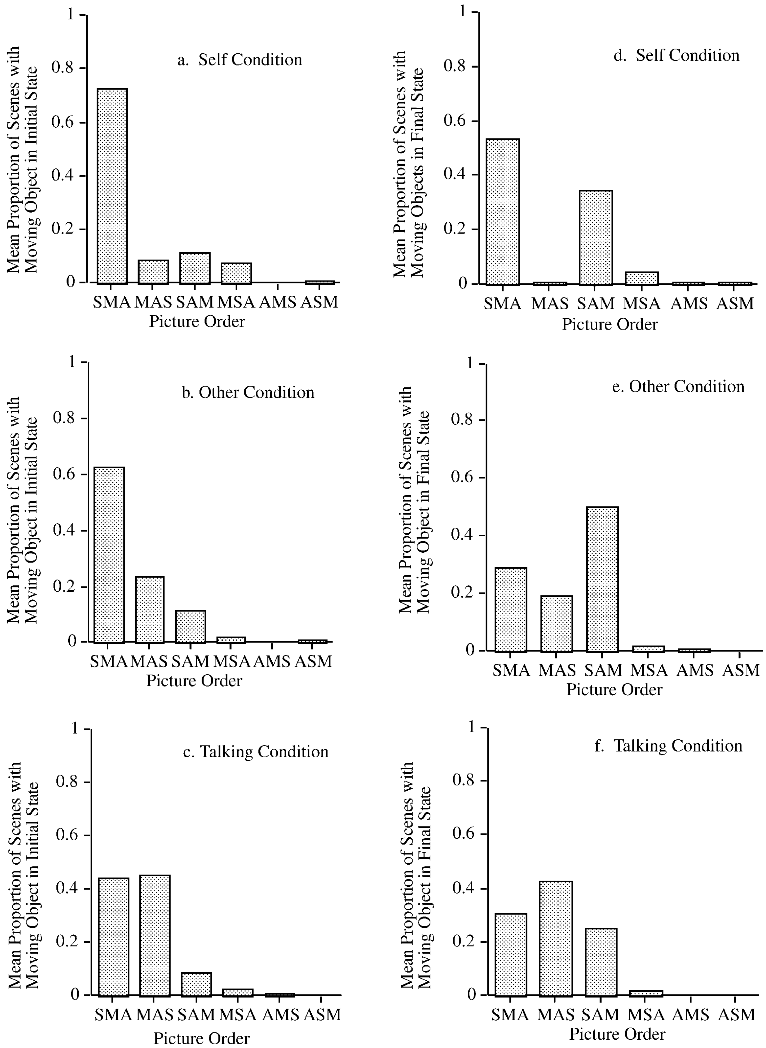

Fig. 5. Mean proportion of scenes that participants reconstructed using an SMA, MAS, SAM, MSA, AMS, or ASM ordering of pictures in the Self (a and d), Other (b and e), and Talking (c and f) conditions. The Moving object was positioned in its *Initial* state in the three graphs on the left side of the figure (a–c), and in its *Final* state in the three graphs on the right side of the figure (d–f). M, Moving object; A, Action; S, Stationary object.

Turning to MAS, we found a robust effect of Condition, $F(2, 57) = 5.53$, $p = .006$, a marginal effect of State, $F(1, 57) = 3.55$, $p = .06$, and no interactions. Planned comparisons revealed that participants in the Talking condition used the MAS ordering significantly more often than participants in both the Self condition ($F(1,57) = 24.83$, $p = .002$), and Other condition ($F(1,57) = 4.08$, $p = .05$).

Finally, for the SAM ordering, we found a main effect of State, $F(1,57) = 24.87, p < .0001$, no effect of Condition, and no interactions. Participants used the SAM ordering significantly more often when the Moving object was in the Final state than when it was in the Initial state.

9. Conclusion

Though they might have, the participants in Study 2 did *not* pick pictures up in random order when asked to use those pictures to reconstruct a scene. Regardless of condition, participants picked up the pictures in a consistent order based on semantic role. Moreover, the particular orders the participants followed differed as a function of condition, and as a function of the state of the Moving object in the pictures.

When the Moving object was positioned where it would be at the beginning of the action, 80% of the participants in the Self condition used the SMA order, picking up the picture of the Stationary object before the Moving object, and the Moving object before the Action (e.g., hoop, girl, and jump)—precisely the order found in the gesture conditions of Study 1. SMA thus appears to be a robust ordering of semantic elements, one that does *not* grow out of the need to share information across minds. SMA is the order of choice even when participants are asked to reconstruct the scene for themselves, suggesting that this particular ordering reflects how the scene itself may be conceptualized.

Not surprisingly, when participants are asked to *talk* while reconstructing the scene (the Talking condition), the order in which they pick up pictures of elements of the scene is influenced by English word order—participants are more likely to pick up the picture of the Moving object before the Action, and the Action before the Stationary object (i.e., canonical English word order, e.g., girl, jump, and hoop). Thus, when using a conventional language (English) during the picture reconstruction task, participants deviate from the SMA order they use when conceptualizing the scene for themselves, and instead adopt an order that more closely resembles the conventional language. Interestingly, when asked to reconstruct the scene so that the reconstruction is transparent to another person (the Other condition), participants also deviate from SMA towards the English order, perhaps because they assume their potential viewers will be English-speakers.

In addition, participants deviate from the SMA order when they are asked to reconstruct an action that has been completed (with the Moving object pictured in its final state rather than its initial state). Participants continue to place the Stationary object in first position—a non-English pattern. However, they switch the order of the Action and the Moving object, placing the

Moving object *after* the Action rather than before it (e.g., hoop, jump, and girl)—again, a non-English pattern, but one that reflects the perceptual properties of the pictured scene (presumably the fact that the object is pictured in the state it would assume *after* the action has already taken place).

Thus, there appear to be two pressures pulling participants away from the SMA order they use when they conceptualize the scene for themselves—canonical English word order, and the perceptual configuration of the scene.

10. General discussion

In Study 1 we found that when asked to communicate to another person using gesture and no words, adults do not take full advantage of the potential for analog representation that the manual modality offers. Rather than represent the scene globally in a single mimetic depiction (much as they would in their spontaneous gestures, cf. Goldin-Meadow et al., 1996; McNeill, 1992), the adults created individual gestures for each semantic role in the scene (the Moving object, the Action, and the Stationary object) and combined those gestures into strings. In other words, they *segmented* the scene into its component parts and then *combined* those parts into a sequenced unit.

10.1. The resilience of segmentation and combination

In previous work, Goldin-Meadow (1982) (see also Goldin-Meadow, 2002) has identified segmentation and combination as "resilient properties" of linguistic systems—properties that appear even when the conditions of language acquisition vary greatly from the conditions children typically experience. Segmentation and combination are found in a child's communications even when the child is lacking a conventional language model (Goldin-Meadow & Mylander, 1990), or when the language model that the child does have is degenerate (Singleton & Newport, 1992), or when the child is first exposed to a language model after puberty, well beyond the age language acquisition typically begins (Curtiss, 1977; Goldin-Meadow, 1978; Newport, 1990). These properties thus appear to be overdetermined in child communication, with a variety of trajectories leading to their development.

The findings of our study indicate that segmentation and combination also appear in the gestures created by hearing adults asked to abandon speech and use only their hands to communicate. These properties thus appear to be essential, not just to child communication, but to adult communication as well. Why are these properties so important to communication? Freyd (1983) has suggested that segmentation and combination are natural outgrowths of the need to share information across minds. She argues that even if knowledge of the physical world is more accurately represented in a continuous form, the process of sharing that knowledge with another individual forces the knowledge through a discrete filter. As a result, its representation ends up looking discrete.

The task that faced the adults in our experiment was to describe a simple scene to another person without using words. The first step in such a task is to make it clear what the scene is about; that is, to establish reference. The need to establish reference leads naturally to breaking up the scene so that each element is identified—it leads to segmentation. For example, although an adult can "jump" a loosely held hand from one point to another to describe a scene in which a girl jumps into a hoop (a gesture that is very likely to be used to accompany a spoken description of this scene, cf. Goldin-Meadow et al., 1996), such a gesture on its own does not indicate who is doing the jumping, nor where the jumping ends up. To indicate who does what to whom, the gesturer needs to produce a series of gestures for the semantic elements of the scene. For example, the gesturer could form curls on his own head to indicate the girl and a C-shaped hand to indicate the hoop, and produce them along with the "jump" gesture.

In our study, the adults produced segmented gestures representing the moving object, the stationary object, and the action. But other semantic elements can be explicitly mentioned in descriptions of motion events (cf. Talmy, 1985, 1991). Although we did not code for these other semantic elements in this analysis, we have coded all of the elements identified by Talmy as core to a motion event in an analysis of the self-styled gesture systems created by deaf children in two cultures (Chinese and American; Zheng & Goldin-Meadow, 2002). Like the hearing adults in our current study, the deaf children in both cultures produced gestures primarily for figures (moving objects), endpoints (stationary objects), and paths (actions). The children only infrequently expressed other elements, such as origin or place. Taken together, these findings suggest not only that the process of segmentation is robust in communication, but that the outcome of this process (i.e., these particular semantic elements) may also be basic to the communication of motion events.

10.2. The naturalness of ordering semantic elements in a combination

Once a scene has been segmented into elements, the elements need to be re-combined for the scene to be adequately described. If left uncombined, the gestures do not convey the relation among the elements within the scene. The most obvious strategy is to conjoin the gestures for these elements into a single gesture string. This is, in fact, the strategy followed by the adults in Study 1; they produced gestures for the elements in a scene without breaking the flow of movement between those gestures.

But merely combining gestures into a single unit is not enough. The gestures for the elements need to be brought together in such a way that the relation among them is clear. The question that we sought to answer in our study was whether the position a gesture occupied within a string would reflect the role played by the element that the gesture represents; that is, whether the adults would follow a consistent *ordering* of semantic roles in their gestures.

We found that the adults did indeed exhibit consistent ordering in their gestures. Importantly, they did so independent of the communicative contexts we varied—that is, whether or not the communication partner was permitted to give feedback; whether or not the information to be communicated was present in the context that partner and gesturer shared; whether or not the gesturer assumed the role of gesture receiver as well as gesture producer. The adults in all five of the conditions in Study 1 followed a consistent order when producing gestures representing each of the semantic roles in the scene. At least 83% of each adult's gesture strings were patterned in precisely the same way, suggesting that *ordering* itself is a natural strategy that adults make use of when they communicate.

The appearance of ordering in the hearing adults' gestures is particularly striking because ordering is not typically found in the gestures hearing adults use on a daily basis (McNeill, 1992). Even more impressive is the fact that ordering appeared so quickly in the adults' gestures and under all of the communicative circumstances we tested. With essentially no time for reflection on what might be fundamental to language-like communication, when asked to use only their hands to communicate under a variety of circumstances, the adults in our study produced gestures characterized by an ordered combination of elements. Thus, ordering also appears to be resilient in all symbolic human communication, not just child systems.

Is ordering a strategy that humans use even when *not* communicating? Study 2 used a non-communicative context to address this question. Adults were asked to reconstruct a scene using pictures. In one of the conditions, the adult was not watched during the reconstruction and believed that we were interested only in the *product* and not the *process* of stacking the pictures. We found that, here again, at least 83% of each adult's reconstructions were created by stacking the pictures according to a single order. It appears that ordering is a natural strategy adults use, not only when communicating with others, but also when conceptualizing a scene for themselves.

10.3. Is there a privileged order in communication and thought?

Not only did the adults in Study 1 use a consistent order when producing gestures for semantic elements, they used the *same* ordering of semantic roles—in all five conditions. They produced gestures for Stationary objects before gestures for Moving objects, and gestures for Moving objects before gestures for Actions (SMA). None of the communicative conditions that we manipulated affected the order that the adults relied on in their gestures.

SMA turns out to be the order used to convey crossing-space events in ASL, and thus may be a natural order for communication in the manual modality. But might SMA have the potential to be more? Might SMA reflect the natural way that humans conceptualize scenes of this sort—whether or not they are communicating about the scenes?

The results from Study 2 support this hypothesis. Adults asked to reconstruct the scenes for themselves consistently picked up pictures of Stationary Objects before pictures of Moving Objects, and pictures of Moving Objects before pictures of Actions. In other words, they followed the SMA order when selecting pictures to reconstruct the scene for themselves. SMA thus appears to reflect the way individuals organize the scene once it has been broken up into components. It appears to be a default ordering, used when adults conceptualize a crossing-space scene for themselves and when they describe the scene to another in the manual modality.

Study 2 attempted to move the phenomenon beyond the communication situation and beyond the manual modality. However, our findings may be limited in two ways—we explore ordering in only one event type (crossing space), and using only visuo-spatial symbols (gestures, pictures). The SMA ordering may therefore be a natural way of conceptualizing a crossing-space scene, but perhaps limited to scenes symbolized in the visual-spatial modality, either in gesture or in pictures (cf. Emmorey, 1996, p. 178). We suggest that even this more limited conclusion is surprising given that visuo-spatial formats are particularly good at supporting simultaneous representations, not sequentially ordered representations. Moreover, the two visuo-spatial symbols we used—gestures and pictures—are distinct representational formats: Pictures are static and two-dimensional, gestures are dynamic and three-dimensional. Thus, at a minimum, SMA reflects a natural conceptualization for two very different visuo-spatial representational systems.

10.4. Pressures that disrupt the default ordering

10.4.1. Conventional language

Although robust, the default SMA order can be overridden. One pressure that can interfere with SMA order is conventional language. Speakers do not adhere to the SMA order when they describe scenes of this sort in English—the canonical English word order for such a scene is MAS (Moving object, Action, Stationary object, "girl jumps into hoop"), not SMA. We found that if English-speaking adults are asked to reconstruct a scene while talking, many not only fail to use the default SMA order in their talk but they also fail to use it in their simultaneously performed reconstructions, organizing them instead around the English MAS order. Even if English-speaking adults are asked to reconstruct the scene without talking but with the goal that "others" can later interpret a videotape of the reconstruction, many will again organize their reconstructions around an MAS rather than an SMA order—presumably because the adults imagine those "others" to be English-speakers.

Thus, the word order dictated within a conventional language can force speakers of that language away from what appears to be a default way of viewing scenes of this sort. Recent findings by Griffin and Bock (2000) support this view. Their data suggest that focusing on the stationary object—or in their scenes, the patient—may be a default way of viewing an action, a perspective that is altered when adults are asked to talk about the scene. Griffin and Bock monitored eye movements under several conditions: Adults described a simple event shown in a picture, with or without the opportunity to prepare their speech, (Speech conditions); adults viewed the picture with the goal of finding the person or thing being acted on in each event (Patient condition); or adults viewed the picture without any specific task requirements (Inspection condition). The interesting result from the perspective of our studies is that the adults' eye movements were skewed toward the patient early in the viewing, not only in the Patient condition, but also in the Inspection condition. Thus, when given no instructions, the adults' first inclination was to focus on the object that is not doing the acting—the semantic element that typically occupies the initial position in the gesture sentences created by the hearing adults in our studies.[1] In contrast, when asked to describe the scene in speech, the adults skewed their eye movements to the object doing the acting, the agent that typically occupies the initial subject position in an English sentence.

If, as it appears, SMA is such a natural way of conceptualizing scenes of this sort, the perplexing question is—why does the SMA order not crop up more often as the default or canonical order in natural languages (Greenberg, 1966; Ruhlen, 1975; Tomlin, 1986)? Slobin (1977) has addressed questions of this sort by identifying a number of pressures that language faces—pressures to be clear, processible, quick and easy, and expressive. Importantly, Slobin points out that these pressures do not necessarily all push language in the same direction. Thus, for example, the pressure to be semantically clear may come into conflict with pressures to be processed quickly or to be rhetorically expressive. The need to be clear may pressure languages to adopt structures that reinforce the SMA order. However, at the same time, the need to be quick and expressive may pressure languages toward structures that do not have this order.

If SMA is indeed as natural a way of conceptualizing this type of scene as our gesture and picture reconstruction data suggest, there may be a cogni-

[1] In the present study, the stationary object was not affected by the action and thus was not a patient. However, two other gesture-creation studies have been conducted in which the stationary object was playing the role of patient (Goldin-Meadow, Yalabik, & Gershkoff-Stowe, 2000; Hammond & Goldin-Meadow, 2002). The hearing adults in these studies tended to produce gestures for the stationary object/patient in initial position in their gesture strings, again going against typical English word order.

tive cost to overriding it, that is, to being forced to view the scene in a different manner. English-speakers may well expend more effort when processing sentences that deviate from this order (i.e., sentences that follow an MAS order) than sentences that adhere to the order (sentences that follow an SMA order). Simultaneously performed tasks that impose a cognitive load on the performer (e.g., Goldin-Meadow, Nusbaum, Garber, & Church, 1993) might be used to test this hypothesis.

10.4.2. Temporal perspective on an event

A second pressure that can interfere with an adult's following the default SMA order is the temporal stance the adult takes to the event in question. Our picture reconstruction task was done under two conditions—the Moving object was pictured in the position it assumed at the beginning of the event, *before* the action had taken place; or the Moving object was pictured in the position it assumed at the end of the event, *after* the action had taken place. This manipulation had a large effect on the orders the adults followed when reconstructing the scene (note, however, that the manipulation had *no* effect on whether the adults adhered to a consistent order—they all ordered the pictures consistently, independent of which order they chose to adopt). When the Moving object was in Initial state, adults reconstructing the pictures for themselves adhered to the natural SMA order. However, when the Moving object was in Final state, the adults began to abandon the default SMA order and organize their reconstructions around an SAM order. Note that the SAM order maintains the Stationary object in first position, but places the Moving object *after* the Action.

The appearance of a reliable SAM order is important for two reasons—methodological and theoretical. Methodologically, it is possible that when the hearing adults placed pictures in the SMA order, they did so only because the diagrammatic action picture was difficult to interpret on its own and only made sense when ordered after all the other pictures. But in the SAM order, the action picture is *not* last, making it clear that the diagrammatic action picture can be interpreted when it appears in other positions, and suggesting that the SMA order is not an artifact of the pictures we used. Theoretically, the fact that two different orders (SMA and SAM) arise as a function of the perspective the participant takes on the scene may offer a cognitive basis for the multiple ordering patterns some languages exhibit. We have found that people organize an event differently as a function of its completeness in a non-communicative task. This cognitive propensity may be one of the pressures that has, over generations, led some languages to develop different constructions to describe an event when it is completed vs. when it is still in progress.

We have identified pressures that pull participants away from the SMA order they use when they conceptualize a scene for themselves. Equally interesting, there appear to be pressures that do *not* pull adults away from the SMA order. None of the manipulations of communicative contexts that we introduced into Study 1 had any effect on the gesturers' ordering of semantic elements. A priori one might think that whether a communication partner provides feedback to gesturers ought to affect the way they present their gestures. However, our findings indicate that, in terms of ordering of symbols for semantic roles, feedback has no effect—nor does shared knowledge, or alternating between roles of gesturer and receiver. Thus, the SMA order is robust against a number of communicative pressures that one might have thought would affect the ordering of symbols for semantic roles.

Of course, we have not exhausted the list of communication pressures that could be manipulated. We showed participants isolated events that could all be described in single simple sentences. We therefore gave the adults no reason to generate a number of sentences that depend on and influence one another within a larger discourse. For example, there was no demand to maintain continuity by referring back to what the gesturer had conveyed in a previous scene, nor any pressure to distinguish new information from old (i.e., a given/new distinction). If these discourse factors had been part of the participants' communication situation, perhaps orders other than SMA would have emerged. The role of connected discourse in shaping ordering patterns could easily be explored using our gesture creation paradigm.

In sum, we have found in two studies—one in which adults were asked to create a "language" to describe a scene to another, and a second in which adults were asked to reconstruct the scene for themselves or for others—that it is natural for humans to sequence symbols that represent semantic roles according to a consistent order. Whether or not adults are communicating, they place symbols for particular semantic roles in particular sequential positions. The reliance on ordering devices found in all natural languages thus appears to reflect general processing strategies that are not specific to language.

However, the *particular* orders that some languages (including English) adopt do not appear to reflect a "natural" sequencing of semantic roles. We have found what appears to be a default order in which symbols for semantic roles are sequenced, at least when those symbols have visuo-spatial qualities. This order is used both when adults conceptualize for themselves a crossing-space event, and when they communicate the event to another in gesture without talking. But this is not the canonical order that most conventional languages offer their speakers to describe such events—conventional language often overrides the default order (whether there is a cognitive cost to doing so is an open question). Thus, although ordering itself is a general cognitive skill that all languages exploit, the particular orders that languages adopt appear to be quite specific to language. They do not necessarily reflect a general (i.e., non-language) way of viewing the world, but instead may be arbitrary outgrowths of the many pressures that conspire to make language what it is.

Acknowledgments

We wish to thank Veronica Sansing for her many valuable contributions to this research. We also thank Carolyn Mylander and the Goldin-Meadow Lab at the University of Chicago for helpful discussion at different stages of the project, and Deborah Kruglak for her assistance in collecting data. This work was supported by NIH Grant No. F 32 DC 00264 and by a Faculty Development grant from Carnegie Mellon University to Lisa Gershkoff-Stowe and by Grant No. RO1 DC00491 from NIH to Susan Goldin-Meadow.

References

Anderson, A. H., & Boyle, E. A. (1994). Forms of introduction in dialogues: Their discourse contexts and communicative consequences. *Language and Cognitive Processes, 9*, 101–122.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

Bock, J. K., & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior, 19*, 467–484.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Clark, H. H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.

Clark, H. H. (1996). *Using language*. London: Cambridge University Press.

Clark, H. H., & Brennan, (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: APA.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472–517.

Clark, H. H., & Schaefer, E. F. (1987). Concealing one's meaning from overhearers. *Journal of Memory and Language, 26*, 209–225.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.

Comrie, B. (1981). *Language universals and linguistic typology*. Chicago: University of Chicago Press.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day "wild child"*. New York: Academic Press.

Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the "model of the listener". In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore Festschrift for Lila Gleitman* (pp. 105–129). New Jersey: Erlbaum.

Emmorey, K. (1996). The confluence of space in language in signed language. In P. Bloom (Ed.), *Language and space* (pp. 171–209). Cambridge, MA:MIT Press.

Feldman, H., Goldin-Meadow, S., & Gleitman, L. (1978). Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock (Ed.), *Action, symbol, and gesture: The emergence of language*. New York: Academic Press.

Fischer, S. A. (1975). Influences on word-order change in American Sign Language. In C. N. Li (Ed.), *Word order and word order change* (pp. 1–25). Austin, TX: University of Texas Press.

Freyd, J. F. (1983). Shareability: The social psychology of epistemology. *Cognitive Science, 7*, 191–210.

Gernsbacher, M. A., & Hargreaves, D. (1992). The privilege of primacy: Experimental data and cognitive explanations. In D. L. Payne (Ed.), *Pragmatics of word order flexibility*. Amsterdam: John Benjamins Publishing Co.

Goldin-Meadow, S. (1978). A study in human capacities. (Review of Genie: A psycholinguistic study of a modern-day 'Wild Child' by S. Curtiss). *Science, 200*, 649–651.

Goldin-Meadow, S. (1982). The resilience of recursion: A study of a communication system developed without a conventional language model. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 51–77). New York: Cambridge University Press.

Goldin-Meadow, S. (2002). The resilience of language. In J. Werker & H. Wellman (Eds.), *Essays in Developmental Psychology Series*. New York: Psychology Press.

Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review, 103*, 34–55.

Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: The non-effects of parental input on language development. *Science, 221*, 372–374.

Goldin-Meadow, S., & Mylander, C. (1984). Gestural communication in deaf children: The effects and non-effects of parental input on early language development. *Monographs of the Society for Research in Child Development, 49*, 1–121.

Goldin-Meadow, S., & Mylander, C. (1990). Beyond the input given: The child's role in the acquisition of language. *Language, 66*(2), 323–355.

Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature, 91*, 279–281.

Goldin-Meadow, S., Nusbaum, H., Garber, P., & Church, R. B. (1993). Transitions in learning: Evidence for simultaneously activated strategies. *Journal of Experimental Psychology: Human Perception and Performance, 19*(1), 92–107.

Goldin-Meadow, S., Yalabik, E., & Gershkoff-Stowe, L. (2000). The resilience of ergative structure in language created by children and by adults. *Proceedings of Boston University conference on language development, 24*, 343–353.

Greenberg, J. H. (1963). *Universals of language. Report of a conference held at Dobbs Ferry, NY*. Cambridge, MA: MIT Press.

Greenberg, J. H. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language*. Cambridge, MA: MIT Press.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*, 274–279.

Hale, K. (1992). Basic word order in two "free word order" languages. In D. L. Payne (Ed.), *Pragmatics of word order flexibility*. Amsterdam: John Benjamins Publishing Co.

Hammond, A. J., & Goldin-Meadow, S. (2002). The robustness of non-English sequences in created gesture systems. In B. Skarabela, S. Fish, & A. H.-J. Do (eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (Vol. 1, pp. 278–289). Somerville, MA: Cascadilla Press.

Hawkins, J. (1983). *Word order universals*. New York: Academic Press.

Hoffmeister, R. (1978). The development of demonstrative pronouns, locatives and personal pronouns in the acquisition of American Sign Language by deaf children of deaf parents. Unpublished Ph.D. Dissertation, University of Minnesota.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60–99.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston, MA: Allyn and Bacon.

Keenen, E. L. (1976). Towards a universal definition of "subject". In C. Li (Ed.), *Subject and topic*. London: Academic Press.

Krauss, R. M. (1987). The role of the listener: Adressee influences on message formulation. *Journal of Language and Social Psychology, 6*, 81–97.

Krauss, R. M., & Fussell, K. (1991). Perspective taking in communication: Representations of others' knowledge in reference. *Social Cognition, 9*, 2–24.

Liddell, S. K. (1980). *American sign language syntax*. The Hague: Mouton.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

Mithun, M. (1992). Is basic word order universal?. In D. L. Payne (Ed.), *Pragmatics of word order flexibility*. Amsterdam: John Benjamins Publishing Co.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11–28.

Newport, E. L., & Ashbrook, E. F. (1977). The emergence of semantic relations in American Sign Language. *Papers and Reports on Child Language Development, 13*, 16–21.

Newport, E. L., & Meier, R. P. (1985). The acquisition of American Sign Language. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition, Vol. 1: The data*. New Jersey: Erlbaum.

Osgood, C. E. (1983). In R. W. Rieber & G. Voyat (Eds.), *Dialogues on the psychology of language and thought*. New York: Plenum Press.

Osgood, C. E., & Bock, J. K. (1977). Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory* (pp. 89–140). New Jersey: Erlbaum.

Osterhout, L. (1994). Event-related brain potentials as tools for comprehending language comprehension. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 15–44). New Jersey: Erlbaum.

Romaine, S. (1992). The evolution of linguistic complexity in pidgin and Creole languages. In J. A. Hawkins & J. Gell-Mann (Eds.), *The evolution of human languages*. Reading, MA: Addison-Wesley.

Ruhlen, M. (1975). *A guide to the languages of the world*. Stanford: Language Universals Project, Stanford University.

Sapir, E. (1921). *Language*. New York: Harcourt, Brace & World.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211–232.

Singleton, J. L., Newport, E. L. (1992). When learners surpass their models: The acquisition of American Sign Language from impoverished input. Unpublished manuscript, University of Illinois, Urbana, Champaign.

Slobin, D. I. (1977). Language change in childhood and in history. In J. Macnamara (Ed.), *Language learning and thought* (pp. 185–214). New York: Academic Press.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition Volume 2: Theoretical issues* (pp. 1157–1256). New Jersey: Erlbaum.

Sridhar, S. N. (1988). *Cognition and sentence production: A cross-linguistic study*. New York: Springer.

Supalla, T. (1982). Structure and acquisition of verbs of motion and location in American Sign Language. Unpublished Ph.D. Dissertation, University of California at San Diego.

Supalla, T., Newport, E. L., Singleton, J. L., Supalla, S., Metlay, D., & Coulter, G. (in press). *Test battery for American Sign Language morphology and syntax*. San Diego, CA: Dawn Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description. Volume III: Grammatical categories and the lexicon* (pp. 57–149). New York: Cambridge University Press.

Talmy, L. (1991). Path to realization: A typology of event conflation. In: *Proceedings of the Berkeley linguistics society* (Vol. 17).

Tomlin, R. S. (1986). *Basic word order: Function principles*. London: Croom Helm.

Wundt, W. (1900/1970). The psychology of the sentence. In A. L. Blumenthal (Ed.), *Language and psychology: Historical aspects of psycholinguistics* (pp. 20–31). New York: Wiley.

Yau, S. (1985). Sociological and cognitive factors in the creation of a sign language by an isolated member of a hearing community. In W. Stokoe & V. Volterra (Eds.), *SLR '83: Sign language research* (pp. 299–306). Silver Spring, MD: Linstok Press.

Zheng, M., Goldin-Meadow, S. (2002). Thought before language: How deaf children express motion events across cultures. Cognition, in press.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Reading, MA: Addison-Wesley.